


Adaptive Human Behavior and Physiology (2019) 5:297–316
<https://doi.org/10.1007/s40750-019-00118-z>

ORIGINAL ARTICLE



Testosterone and Fathers' Parenting Unraveled: Links with the Quantity and Quality of Father-Child Interactions

Else E. de Vries¹  · Lotte D. van der Pol² · Harriet J. Vermeer¹ · Marleen G. Groeneveld² · Tom Fiers³ · Judi Mesman¹

Received: 31 July 2019 / Accepted: 3 September 2019/

Published online: 2 October 2019

© The Author(s) 2019

Abstract

Objective Individual differences in quality of father involvement in caregiving might in part be explained by fathers' testosterone (T) levels. We examined the links between fathers' ($n = 32$) salivary T levels, amount of time spent with their child (12–30 months of age), type of father-child interaction, and fathers' sensitivity.

Methods During two home visits, video observations of father-child interactions were conducted to measure fathers' sensitivity during a challenging and harmonious interaction. Fathers' saliva was collected several times throughout the day on a working day and on the home visit days, including right before and after each father-child interaction.

Results Fathers' T secretion throughout the day was lower on home visit days (i.e., days with a higher amount of time spent with their child) than on a working day. For both challenging and harmonious father-child interactions, mean T levels did not differ before and after father-child interactions. However, individual changes in fathers' T levels during the father-child interactions did predict fathers' sensitivity. Specifically, the more T increased during the challenging interaction, or decreased during the harmonious interaction, the more sensitive the father was during that interaction as well as during a subsequent interaction.

Conclusions Parenting quality is most optimal when fathers' T system reacts in the expected direction given the context of the father-child interaction, i.e., a T decrease during a harmonious interaction and a T increase during a challenging interaction. Our study underscores the importance of examining the interplay between biology, behavior, and caregiving context in fathers' parenting.

Keywords Fathers · Testosterone · Sensitivity · Parenting · Observation

✉ Else E. de Vries
e.e.de.vries@fsw.leidenuniv.nl

Introduction

Father involvement in childcare is widely recognized as being beneficial to children's social and emotional development (Lamb 2010). Father involvement in caregiving can be conceptualized in many different ways, and often the distinction is made between *quantitative* aspects (e.g., amount of time with child) and *qualitative* aspects (e.g., social-emotional quality of father-child interactions) (Lamb 2000; Pleck 2010). Individual differences in father involvement have an effect on, and are affected by, the hormone testosterone (T). Research shows that having children and investing more time in childcare is in general related to lower T in fathers (e.g., Berg and Wynne-Edwards 2001; Gettler et al. 2011; Kuzawa et al. 2009). Moreover, it seems that lower T levels are associated with more sensitive, responsive, and nurturing behaviors (e.g., Kuo et al. 2015; van Anders et al. 2012; Weisman et al. 2014). However, not all parenting cues elicit a decrease in men's T levels. Simulation studies indicate that challenging father-child interactions increase T, whereas harmonious father-child interactions decrease T (e.g., Fleming et al. 2002; van Anders et al. 2012). To our knowledge, fathers' T levels have not yet been studied in relation to their sensitivity during different types of real-life father-child interactions. In the current study we examined the associations between fathers' T levels and the quantity and quality of father involvement.

Testosterone and the Amount of Time with Child

Both theory and research indicate that the quantity of father involvement can be linked to the hormone testosterone (T). According to the Challenge Hypothesis, decreased T facilitates parenting behavior of men (Wingfield et al. 1990). More specifically, the Challenge Hypothesis proposes that the association between parenting behavior and T is bidirectional: high T inhibits parenting behavior, while at the same time T is downregulated by cues that are related to children or childcare. Fathers have lower basal T levels than childless men, indicating that having children as such may downregulate their T, and lower T levels are related to greater childcare involvement of fathers (Alvergne et al. 2009; Kuzawa et al. 2009; Perini et al. 2012). Even expectant fathers show a decline in T levels before the birth of their child, and greater declines in prenatal T are associated with greater contributions to childcare tasks after birth (Edelstein et al. 2017). A decline in T levels has been found across the transition to parenthood (Berg and Wynne-Edwards 2002; Storey et al. 2000), and in response to an increase in everyday childcare involvement (Gettler et al. 2015). In line with this notion, the physical proximity of a child influences fathers' T levels. Compared with fathers who sleep separately from their child, fathers who sleep on the same surface as their child have lower T levels in the evening and a greater diurnal decline in T (Gettler et al. 2012; Lawson et al. 2017). In sum, many studies indicate that fathers' T system reacts to a higher amount of time spent with their child by lowering T levels or that fathers with lower T levels initiate more contact with their child.

The previously mentioned studies examined the general amount of childcare involvement in relation to T levels on an arbitrarily chosen day. It is also possible to relate the amount of time fathers spend with their child on a specific day to T levels on the same day. In this way, it can be examined whether T fluctuates on a day-to-day basis depending on the daily amount of time spent with the child. However, previous studies

found no differences between T levels on a ‘with-child’ and a ‘without-child’ day (Storey et al. 2011; Gray et al. 2004), suggesting that T levels are independent of the day-to-day amount of time fathers spent with their child. However, this result may be due to measuring T levels on a very small part of the day. Because T levels follow a diurnal rhythm that is not necessarily linear, collecting saliva several times a day is a more thorough approach to capture the individual variability of T, and this approach makes it possible to calculate the overall T secretion throughout the day. To examine whether fathers’ T secretion throughout the day fluctuates depending on the amount of time fathers spend with their child on those days, in our study fathers collected saliva samples several times a day for three days.

Testosterone and Different Types of Father-Child Interactions

Apart from the amount of time with child as a broad measure of parenting, it is also possible to explore parenting more in depth by examining specific parenting situations. Different types of parenting situations may have different effects on T. For example, in a simulated challenging parenting situation (exposure to a crying baby doll), men’s T levels have been found to increase rather than to decrease (Fleming et al. 2002; Storey et al. 2000; van Anders et al. 2012), which is not in accordance with the Challenge Hypothesis that posits that *all* parenting cues should decrease T (Wingfield et al. 1990). To understand these opposing associations between T and parenting situations, van Anders et al. (2011) propose an alternative theory called the Steroid/Peptide Theory of Social Bonds (S/P Theory; van Anders et al. 2011). This theory holds that parent-child contexts that are harmonious (like playing together) or involve nurturance (such as comforting or feeding) will decrease T, while parent-child contexts that involve a challenge (such as managing conflicts or dealing with a distressed child) will increase T. The inconsistent findings thus far regarding the impact of father-child interaction on T (some studies did find an effect, others did not) offer support for the notion that it is important to distinguish between different types of interactions.

Testosterone and Paternal Sensitivity

Prior studies of father involvement and T have generally focused on quantitative aspects such as amount of father involvement and type of father-child interactions. As Gettler (2016) mentions, the question remains whether T relates to *how* fathers behave as a parent. For example, is T related to how sensitive, nurturing, and responsive fathers are towards their child, and is this relation different depending on the type of father-child interaction? It is precisely this *qualitative* aspect of father involvement that is important for positive child development (Pleck 2010). Previous research has shown that lower T levels in fathers are related to more sensitive parenting (Fleming et al. 2002; Storey et al. 2000; van Anders et al. 2012). During laboratory-based father-baby interactions, fathers with a lower baseline T and a higher T decline showed more optimal parenting behaviors such as a higher frequency and duration of affectionate touch, following the gaze of their baby, and baby-directed speech (Weisman et al. 2014). Moreover, a higher diurnal decline in T levels has been associated with more paternal sensitivity and respect for the child’s autonomy during free play (Endendijk et al. 2016).

To our knowledge, only one study included fathers' T levels before and after a father-child interaction as well as fathers' sensitivity during the interaction. When fathers experienced a greater decline in T levels while they comforted their distressed child, they were more sensitive towards their child during a subsequent challenging task (Kuo et al. 2015). This study, however, measured T levels during the first father-child interaction, and sensitivity during the second father-child interaction. It was thus not possible to examine whether change in T levels during a father-child interaction was related to fathers' behaviors within one and the same father-child interaction. Moreover, Kuo et al. (2015) were not able to examine the direction of the association between fathers' T levels and sensitivity.

The Current Study

The aim of the current study was to examine the links between fathers' salivary T levels, the amount of time with child, different types of interaction between father and child, and fathers' level of sensitivity. Because T levels of men follow a diurnal rhythm with highest T levels in the morning, steeply declining T levels before noon, a slower decline in the afternoon and early evening, reaching lowest T levels in the evening (Booth et al. 2006; Matsui et al. 2009), it is necessary to measure T several times throughout the day. Therefore, fathers' saliva was collected several times throughout the day on two home visit days and a working day, which enabled us to calculate fathers' overall T secretion throughout these days. Fathers spent a higher amount of time with their child on the home visit days than on the working day. During both home visits, father and child were involved in two types of structured observation episodes to elicit a harmonious and challenging interaction. A discipline episode was chosen to elicit a challenging interaction between father and child, whereas a play episode was intended to elicit a harmonious interaction. Fathers' T-responses to these different types of interaction (i.e., challenging and harmonious) were measured, and video-observations were made of the father-child interactions to measure fathers' sensitivity.

Based on the literature discussed above, we tested three hypotheses. First, we expected that fathers' T secretion throughout the day would be *lower* on the home visit days (i.e., days with a higher amount of time spent with their child) than on a working day (1). Second, we expected that T would *increase* after a challenging father-child interaction (2A), and T would *decrease* after a harmonious father-child interaction (2B). Third, we expected that a higher T *increase* during a challenging father-child interaction would be related to lower sensitivity (3A), and a higher T *decrease* during a harmonious father-child interaction would be related to higher sensitivity (3B). Finally, a more exploratory aim was to examine the direction of this association: do fathers' T levels predict fathers' sensitivity, or does fathers' sensitivity predict fathers' T levels?

Methods

Participants

Fathers with a child between 12 months and 30 months of age at the time of recruitment were eligible for participation. Exclusion criteria were severe physical or intellectual

impairments of parent or child, and fathers not speaking the Dutch language. Fathers were asked to participate in two home visits. In addition to the home observations, participation in the study included collecting saliva samples multiple times, computer testing, and filling in questionnaires. Between March and June 2015, eligible fathers were recruited in the Western region of the Netherlands by distributing a brochure with information about the study at several public locations known to attract parents of young children such as day care centers, swimming pools, and parenting meetings. Additionally, a Facebook page was created with information about the study and contact details of the research team. Fathers were also recruited within the authors' network and by contacting fathers who previously participated in an observational study conducted by our research team.

Recruitment resulted in 32 participating fathers who were aged between 23.4 and 51.8 years ($M = 38.1$, $SD = 5.8$). All fathers, except one, lived with the biological mother of the child. With regard to educational level, most fathers had finished academic or higher vocational schooling (69%). The children were on average 2.1 years old ($SD = 0.6$). Fifty-three percent of the children were girls. The majority of the children had one or more siblings (63%).

Procedure

Fathers were visited twice. During both home visits, father and child were involved in two types of structured observation episodes to elicit a harmonious and challenging interaction. We will refer to these as the play episode and discipline episode respectively. During the play episode, father and child received a bag with toys and were free to play with the toys for 15 min. Fathers were instructed to play with their child the way they would normally do. During the discipline episode, fathers received another bag with attractive toys and were given the instruction to take the toys out of the bag but to not let their child touch the toys for two minutes. Subsequently, the child was only allowed to play with the least attractive toy (a stuffed animal) for another three minutes, after which the task was finished. In the 'Play-Discipline visit' fathers first interacted with their child during the play episode and subsequently during the discipline episode. In the 'Discipline-Play visit' the order of the episodes was reversed. The duration between the two observation episodes was 15 (Discipline-Play visit) to 35 minutes (Play-Discipline visit). The home visits were planned about two weeks apart and the order of the home visits was counterbalanced between fathers. During the home visits father-child interactions were filmed, and fathers completed a computer test. They also filled in a set of questionnaires to provide information on sociodemographic factors, psychological complaints, family life (e.g., role division), as well as information on factors potentially associated with hormone levels (e.g., weight, medication, and physical activity).

To measure fathers' T levels, they were asked to collect saliva samples by passive drool in polypropylene tubes on the two visit days and on a working day between the two home visits, which will be referred to as the reference day. See Fig. 1 for an overview of the structure of each visit and all moments of saliva collection. The T measurements on the reference day were meant to capture the T pattern on a working day during which the father had little contact with his child. On all three days, fathers were instructed to collect saliva directly after waking in the morning and right before going to bed in the evening. In addition, saliva was collected at several time points in the afternoon (see Fig. 1). In total, each father collected 15 saliva samples: six samples

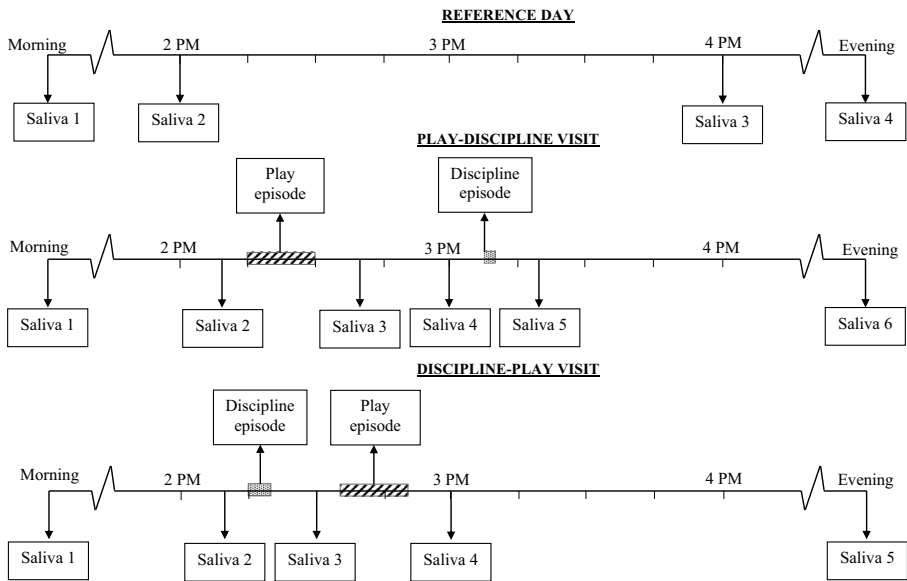


Fig. 1 Overview of the reference day, home visits, and moments of saliva collection *Note.* During the Play-Discipline visit, the extra saliva collection after the play episode (“Saliva 4”) was meant as a baseline measure with which we were able to compare the T level after the discipline episode (“Saliva 5”)

on the day of the Play-Discipline visit, five samples on the day of the Discipline-Play visit, and four samples on the reference day.¹ During the home visits, which roughly took place between 2 PM and 4 PM, saliva was collected right before and 10 to 15 min after each observation episode. Because salivary changes in T are detectable approximately 10 to 15 min after a stimulus or interaction (e.g., Roney et al. 2007; van Anders et al. 2014), fathers were asked to rest (i.e., to not interact with their child) for 10 to 15 min during the period between each observation episode and subsequent saliva collection. On the reference day, fathers collected saliva at 2 PM and 4 PM to resemble the start and end times of the visit days. Saliva samples were stored in the parent’s freezer until pick-up by the research team and were then stored at -80°C until analysis.

All home visits were conducted by a trained (under)graduate student, or the first or second author. After each visit the child received a small present, and after the second visit the father received a gift of 30 euros. Informed consent was obtained from all participating families. Ethical approval for this study was obtained from the Commission Research Ethics Code of the Leiden Institute of Education and Child Studies.

Measures

Fathers’ T Levels Salivary samples were analyzed at the endocrinology laboratory at Ghent University Hospital (Belgium). After centrifugation at 2000 g for five minutes,

¹ Note that during the Play-Discipline visit, we collected saliva twice after the play episode, which enabled us to examine (a) whether (and when) T returned to a baseline level after the hypothesized decrease in T due to a harmonious parent-child interaction, and (b) whether T would subsequently increase after a challenging parent-child interaction in comparison with the baseline level. During the Discipline-Play visit, we collected saliva only once after the challenging episode to examine T changes as a result of different types of consecutive parent-child interactions.

250 μL supernatant from passive drooling was collected and kept at -80°C until LC–MS/MS analysis. A liquid–liquid extraction (LLE) was used: 20 μL d3-T internal standard was added before extraction with 2 mL of diethylether; after mixing for three minutes samples were frozen and decanted with subsequent drying of the collected supernatant; the dried supernatant was then reconstituted in a final solution of 125 μL methanol of which 100 μL was injected for liquid chromatography. T was acquired from Sigma–Aldrich (Saint Louis, USA), and d3-T from CDN Isotopes (Quebec, Canada). All standards and internal standards were dissolved in methanol. Methanol, water and acetonitrile were LC–MS grade from BioSolve BV (Varkenswaard, The Netherlands).

For measurement of salivary T by LC–MS/MS an AB Sciex 5500 triple-quadrupole mass spectrometer (AB Sciex; Toronto, Canada) was used, coupled with an APCI probe on the Turbo-V source. The liquid chromatography system consisted of a Shimadzu system using a C8 security guard column (5 μm , 4×2 mm) and a C8 Luna analytical column (3 μm , 50×3 mm) (Phenomenex; Torrance, USA). Measurements were performed by the tandem mass spectrometer running in multiple reaction monitoring (MRM) mode by using transitions m/z 289/109/97 for T and 292/109/97 for d3-T. A declustering potential (DP) of 100 V and a collision energy (CE) of 32 eV was used for all the analytes. Data processing was performed through MultiQuant version 2.0.2. For analysis on 250 μL saliva, inter-assay CV is 8.2% at 0.23 ng/dL (8 pmol/L) with an LOQ of 0.07 ng/dL (2.4 pmol/L).

All T values of fathers (ranging from 1.26–15.48 ng/dL), including T patterns throughout the day, were in accordance with values reported in the literature. There were five fathers with one missing T value, and one father with three missing T values. Missing T values were predicted with linear regression using the T variable from another time point that correlated highest with the T variable with missing values. To compare the T secretion throughout the day on the home visit days with a working day (Hypothesis 1), T's area under the curve with respect to the ground (AUCg) was calculated across the morning and evening T measurements for each day separately, using the trapezoid formula (Pruessner et al. 2003). The higher the AUCg value, the more T secretion throughout the day. To examine T change during the different types of father-child interactions (Hypothesis 3), we followed the method of Endendijk et al. (2016) for the calculation of episode-specific T reactivity: $((T \text{ before episode} - T \text{ after episode}) / T \text{ before episode}) * -1$. A positive value of T reactivity represents an increase in T during the episode, whereas a negative value of T reactivity represents a decrease in T. The T reactivity on the reference day was calculated as follows: $((T \text{ 2 PM} - T \text{ 4 PM}) / T \text{ 2 PM}) * -1$.

Fathers' Sensitivity During each home visit, fathers' sensitivity toward their child was assessed during a harmonious father-child interaction (i.e., play episode) and during a more challenging father-child interaction (i.e., discipline episode).

Fathers' sensitivity toward their child during the play episode was coded using an adapted version of the sensitivity and nonintrusiveness scales of the fourth edition of the Emotional Availability Scales (EAS; Biringen 2008; Hallers-Haalboom et al. 2014). Sensitivity in the EAS refers to appropriate responding to the child's signals combined with positive affect. Nonintrusiveness refers to following the child's lead and waiting for optimal breaks to enter interaction without interfering with the child's flow of activities, which in essence is also a form of sensitivity. Each scale consists of seven subscales; the first two subscales are coded on a 7-point Likert-type scale and the other five subscales are coded on a 3-point Likert-type scale. Fathers' sensitivity during the play episodes was

rated by six coders. The two play episodes within the same family were coded by different coders to guarantee independency among ratings. All coders completed a reliability set ($n = 60$). Intercode reliability was adequate with intraclass correlations (single measure, absolute agreement) higher than .70. Given that the nonintrusiveness and sensitivity scales were strongly correlated (Play-Discipline visit: $r = .54$, $p < .01$; Discipline-Play visit: $r = .58$, $p < .01$), a combined standardized mean score was computed per visit to reflect fathers' overall level of sensitivity during the play episode.

Fathers' sensitivity toward their child during the discipline episode was rated using the Erickson scale of supportive presence (Egeland et al. 1990; Erickson et al. 1985). The scale refers to the father's appropriate expression of positive regard and emotional support in response to the child's signals by acknowledging the child's accomplishments, as well as by encouraging, reassuring, calming, or giving a physical sense of support. Supportive presence is coded on a 7-point scale. Fathers' sensitivity during the discipline episodes was rated by two coders. Each coder only rated one discipline episode within a family to guarantee independence among ratings. Furthermore, sensitivity during the discipline episodes was rated by coders who had not coded sensitivity during the play episodes. Intercode reliability was adequate, with intraclass correlations (single measure, absolute agreement, $n = 20$) higher than .70.

Amount of Father-Child Contact after the Observation Episodes Although we asked fathers not to interact with their child after each observation episode (i.e., the resting periods), some fathers did have contact with their child. Given that contact after the observation episode could influence subsequent T levels, the amount of verbal and physical contact between father and child after each observation episode was coded by six trained coders. To guarantee independence among ratings, the two visits were coded by different coders who had not coded fathers' sensitivity. Intercode reliability was adequate, with intraclass correlations (single measure, absolute agreement, $n = 20$) higher than .70.

Data Analysis Plan

All variables were inspected for possible outliers, defined as values more than 3.29 *SD* above or below the mean (Tabachnick and Fidell 2012). Outliers were winsorized by giving them a marginally higher value than the most extreme not outlying value. All variables were normally distributed. We adjusted the analyses for several confounders by adding them as covariates and/or residualizing the T measurements for the confounders² (see below for a more detailed description). Prior to our main analyses, Pearson correlation coefficients were computed between the main study variables.

To examine the first hypothesis, whether the T secretion throughout the day (AUCg) was lower on the home visit days (i.e., a day with a higher amount of time spent with the child) than on a working day, a repeated-measures analysis of variance (ANOVA) was performed with type of day as within-subject factor. Only fathers who worked on the reference day ($n = 26$) were included in this analysis. The analysis was adjusted for fathers' age and weight, which have been robustly related to men's T in the literature, by adding these variables as covariates.

² To minimize the number of covariates in our analyses, we residualized the T measurements for confounders if possible.

For the second hypothesis, the T levels during the home visits were used to test whether (A) T *increased* after a challenging father-child interaction (i.e., discipline episode), and (B) T *decreased* after a harmonious father-child interaction (i.e., play episode). Two repeated-measures ANOVAs (one for each home visit) were conducted with time of saliva collection as within-subject factor. Because contact after each structured observation episode potentially influenced subsequent T measurements, we controlled for the average amount of verbal and physical contact between father and child after the observation episodes by adding this variable as a covariate. Additionally, these analyses were controlled for fathers' age and weight by adding these variables as covariates.

Four linear regression analyses were conducted to test the third hypothesis that (A) a higher T increase during a challenging father-child interaction would be related to *lower* sensitivity, and (B) a higher T decrease during a harmonious father-child interaction would be related to *higher* sensitivity. In these analyses, the episode-specific T reactivity was included as predictor and sensitivity within the same episode as outcome. All episode-specific T reactivity scores were residualized for age, weight, and amount of contact after the interaction episode.

With a more exploratory aim, four linear regression analyses were conducted to examine the direction of the association between fathers' sensitivity and T. In the first set of two analyses, the episode-specific T reactivity during the first episode was included as predictor and sensitivity during the second episode as outcome. In the second set of analyses, sensitivity during the first episode was included as predictor and the episode-specific T reactivity during the second episode as outcome. Again, the episode-specific T reactivity scores were residualized for age, weight, and amount of contact after the interaction episode.

Results

Preliminary Analyses

The means, standard deviations, and bivariate correlations for the main study variables regarding all 32 participating fathers are presented in Table 1. During the Play-Discipline visit, lower T reactivity (i.e., higher T decrease) during the play episode was related to higher sensitivity during the discipline episode. During the Discipline-Play visit, lower T reactivity during the play episode was related to higher sensitivity during this play episode. Higher T reactivity (i.e., higher T increase) during the discipline episode of this visit was related to higher sensitivity during the play episode. Further, the T secretions throughout the day (AUCg) on the three days were highly correlated with each other. With regard to the T levels in each of the home visits separately, the correlations between the T levels were all high during the Play-Discipline visit ($r = .61$ to $.85$, $ps < .001$) as well as during the Discipline-Play visit ($r = .89$ to $.94$, $ps < .001$). During the reference day, the T levels at 2 PM and 4 PM were also highly correlated ($r = .69$, $p < .001$).

Testosterone and the Amount of Time with Child

For the 26 fathers who worked on the reference day (and thus had little contact with their child during that day), we found a near-significant effect of type of day on T

Table 1 Summary of means, standard deviations, and correlations for main study variables ($n = 32$)

	1	2	3	4	5	6	7	8	9	10	<i>M</i>	<i>SD</i>
Play-Discipline visit												
1. Sensitivity play episode											21.02	3.23
2. Sensitivity discipline episode	-.01										4.59	1.68
3. T reactivity play episode	.02	-.48**									.001	0.22
4. T reactivity discipline episode	-.09	.34	-.49**								.05	0.26
Discipline-Play visit												
5. Sensitivity play episode	.39*	-.05	-.01	.02							21.80	2.99
6. Sensitivity discipline episode	-.02	.51**	-.12	.13	.04						4.45	1.64
7. T reactivity play episode	.03	.01	.08	.01	-.35*	-.14					.04	0.17
8. T reactivity discipline episode	.19	-.12	.13	-.27	.37*	-.03	-.41*				.02	0.13
T secretion throughout the day (AUCg)												
9. Play-Discipline visit	-.18	.10	-.24	.16	-.09	-.13	.29	-.22			4020.47	1201.47
10. Discipline-Play visit	.00	.16	-.19	.08	-.01	-.01	.24	-.34	.75***		4001.40	1156.94
11. Reference day	.09	-.05	-.32	.22	-.06	-.12	.22	.03	.61***	.54**	4314.99	1067.11

Means and standard deviations for T represent non-residualized scores. The proportion to which T increased or decreased (ng/dL) during the interaction episode, whereas the calculation of T secretion throughout the day is based on the area under the curve using the raw T levels (ng/dL) and the time distance between each T measurement (minutes). Correlations with T are based on residualized scores. The mean and standard deviation for sensitivity during the play episode are based on fathers' unstandardized mean scores of the sensitivity and nonintrusiveness scales. Correlations with sensitivity during the play episode are based on the standardized mean score of the sensitivity and nonintrusiveness scales

* $p < .05$. ** $p < .01$. *** $p < .001$

secretion throughout the day, $F(2,46) = 3.14$, $p = .053$, $\eta_p^2 = .12$. Contrasts revealed that the T secretion throughout the day during the Discipline-Play visit ($M = 3857.32$, $SD = 1070.49$) was significantly lower than on the reference day ($M = 4293.67$, $SD = 1040.25$), $F(1,23) = 4.54$, $p = .04$, $\eta_p^2 = .17$. Similarly, we found a trend that the T secretion throughout the day during the Play-Discipline visit ($M = 3881.03$, $SD = 1133.18$) was lower than on the reference day, $F(1,23) = 3.89$, $p = .06$, $\eta_p^2 = .15$. The T secretion throughout the day during the Play-Discipline visit did not differ from the T secretion during the Discipline-Play visit, $F(1,23) = .02$, $p = .88$, $\eta_p^2 = .00$.

Testosterone and Different Types of Father-Child Interactions

Figure 2 shows fathers' T levels during the Play-Discipline visit, Discipline-Play visit, and the reference day. No differences were found between fathers' T levels at any time point during the Play-Discipline visit ($F(3,84) = .35$, $p = .79$, $\eta_p^2 = .01$), or the Discipline-Play visit ($F(2,56) = .84$, $p = .44$, $\eta_p^2 = .03$). Thus, T levels before the play episodes did not differ from T levels after the play episodes. Similarly, T levels before the discipline episodes did not differ from T levels after the discipline episodes.

Testosterone and Paternal Sensitivity

For each home visit, we examined whether fathers' T reactivity during the observation episodes was associated with fathers' sensitivity. We first looked at associations between fathers' T reactivity and sensitivity within the same observation episode. With regard to the first episode of each visit, no associations between fathers' T reactivity and sensitivity were found (i.e., the play episode during the Play-Discipline visit and the discipline episode during the Discipline-Play visit). Concerning the second episode, the T reactivity during the discipline episode (in the Play-Discipline visit) was marginally positively linked to fathers' sensitivity during this episode ($B = 2.29$, $SE(B) = 1.17$, $\beta = .34$, $p = .06$; Fig. 3). In other words, contrary to Hypothesis 3A, there was a trend that the more T increased during the discipline episode, the more sensitive the fathers

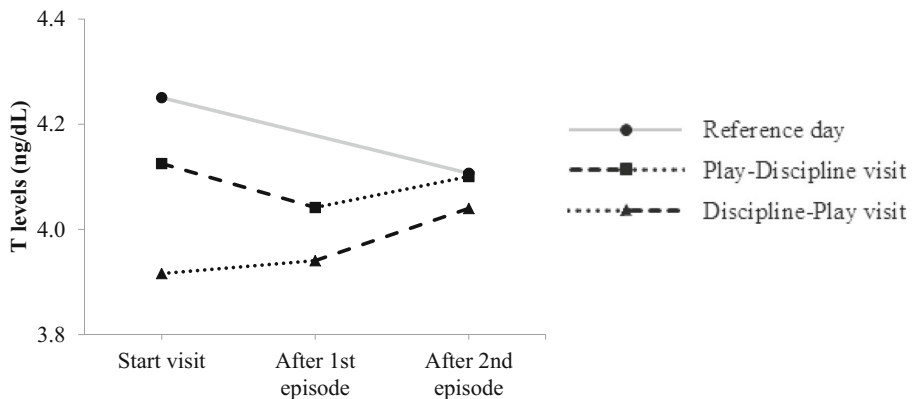


Fig. 2 Fathers' mean T levels (ng/dL) during the Play-Discipline visit and Discipline-Play visit, and the Reference day. *Note.* Fathers' T levels are based on non-residualized scores. The T levels on the reference day were only measured twice to resemble the start and end times of the visit days

were during this episode. Consistent with Hypothesis 3B, the more T decreased during the play episode (in the Discipline-Play visit), the more sensitive the father was during this episode ($B = -1.91$, $SE(B) = .93$, $\beta = -.35$, $p = .049$; Fig. 3).

With a more exploratory aim, four regression analyses were conducted to examine the direction of the association between fathers' sensitivity and T. We found that during the Play-Discipline visit, a higher T decrease during the play episode predicted higher sensitivity during the subsequent discipline episode ($B = 2.29$, $SE(B) = 1.36$, $\beta = -.48$,

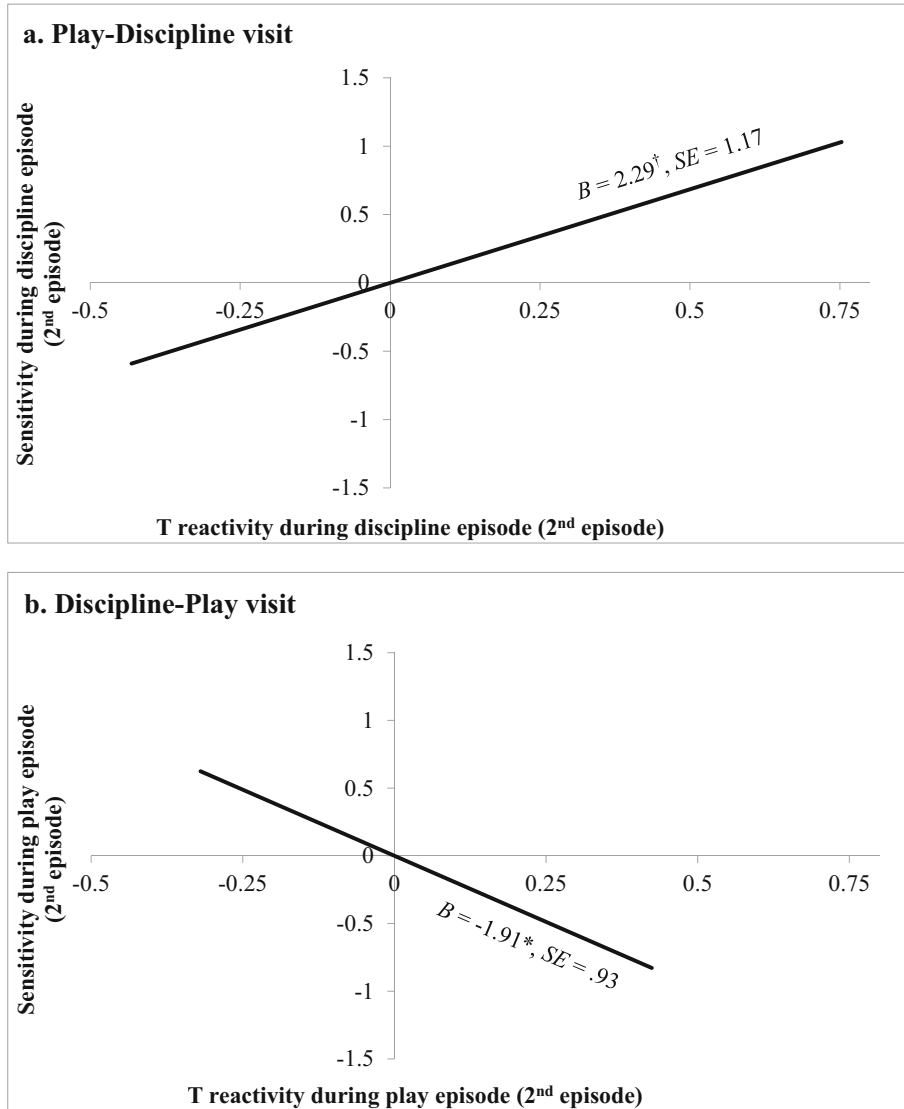


Fig. 3 Associations between fathers' T reactivity and sensitivity within the second observation episode, for the Play-Discipline visit (a) and the Discipline-Play visit (b). *Note.* The T reactivity scores are residualized for age, weight, and contact after the father-child observation episode. The sensitivity score during the play episode is a standardized mean score of the sensitivity and nonintrusiveness scales. The sensitivity score during the discipline episode is a standardized score of the supportive presence scale. * $p < .05$, $\dagger < .10$

$p = .006$; Fig. 4). During the Discipline-Play visit, a higher T increase during the discipline episode predicted higher sensitivity during the subsequent play episode ($B = 2.84$, $SE(B) = 1.29$, $\beta = .37$, $p = .04$; Fig. 4). We did not find associations between fathers' sensitivity during the first observation episode and fathers' T reactivity during the second observation episode in either visit. These findings suggest that fathers' T fluctuations affect fathers' sensitivity rather than the other way around.

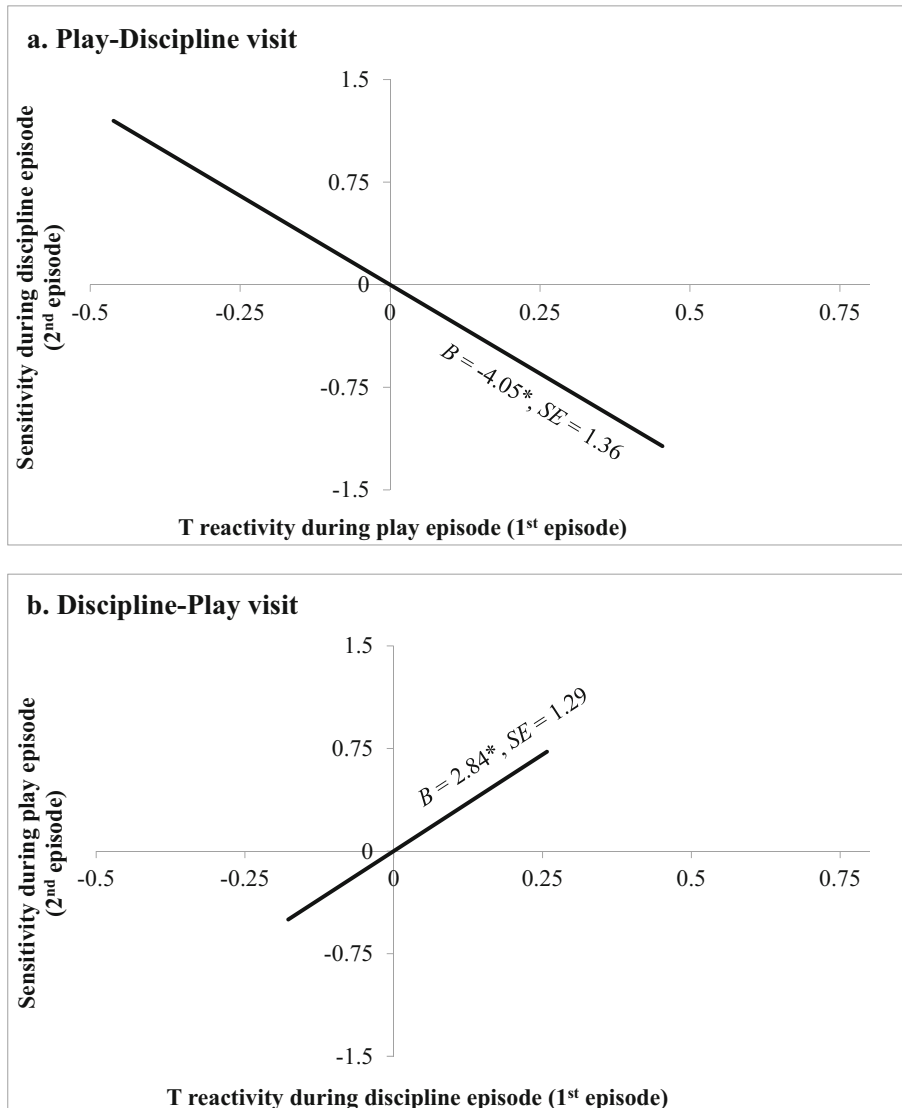


Fig. 4 Associations between fathers' T reactivity during the first father-child observation episode and sensitivity during the second father-child observation episode, for the Play-Discipline visit (a) and the Discipline-Play visit (b). *Note.* The T reactivity scores are residualized for age, weight, and contact after the father-child observation episode. The sensitivity score during the play episode is a standardized mean score of the sensitivity and nonintrusiveness scales. The sensitivity score during the discipline episode is a standardized score of the supportive presence scale. * $p < .05$

Further, we tested whether T reactivity was specifically linked to the father-child interactions and not a result of normal fluctuations of T levels in the afternoon. All results were essentially unchanged when we additionally controlled for T levels on the reference day. Lastly, the associations between fathers' T reactivity and sensitivity were not moderated by visit order.

Discussion

In the current study we explored the links between fathers' T levels, the amount of time with child (i.e., home visit days vs. working day), different types of interaction between father and child (i.e., challenging and harmonious interactions), and fathers' level of sensitivity. In line with our hypothesis, we found that fathers' T secretion throughout the day (AUCg) was lower on days when fathers spent more time with their children (i.e., the home visit days) than on a working day. With regard to type of interactions, for both challenging father-child interactions (i.e., discipline episodes) and harmonious father-child interactions (i.e., play episodes), our hypothesis was not confirmed given that average T levels before the interactions did not differ from average T levels after the interactions. We also examined whether T changes during different types of father-child interactions were related to how sensitive fathers were towards their child. Although we expected decreases in T to be related to higher sensitivity in all father-child interactions, the results indicated that the more T *increased* during the challenging interaction, or *decreased* during the harmonious interaction, the more sensitive the father was during that interaction. Lastly, we explored the direction of the association between fathers' T changes and their sensitivity. Changes in T levels during the first father-child interaction predicted fathers' sensitivity during the subsequent interaction, rather than the other way around.

Testosterone and the Amount of Time with Child

Our findings suggest that the overall T secretion throughout the day (AUCg) fluctuates on a day-to-day basis depending on the amount of time fathers spend with their child on those days. More specifically, fathers' T secretion throughout the day seemed to be lower on days in which they spent more time with their child than on a working day. It should be noted that on one of the two home visit days, there was only a trend that the T secretion was different from the working day. These findings are in agreement with the Challenge Hypothesis (Wingfield et al. 1990) and many previous studies (e.g., Alvergne et al. 2009) that related a lower T secretion to a higher amount of time fathers spent with their children. By measuring the overall T secretion throughout the day on different types of days, our results add to studies that measured T on only a small part of the day or studies that examined how much T changed from morning to night without including time points in between. These previous studies did not capture the nonlinear diurnal rhythm in T. Moreover, due to the computation of the AUCg, the statistical power is increased by combining

information from repeated measurements (Pruessner et al. 2003). By using this thorough approach, our study offers support for the notion that paternal T secretion is to some degree responsive to the amount of time spent with the child or that fathers with lower T secretion spend more time with their child.

Testosterone and Different Types of Father-Child Interactions

Based on the Steroid/Peptide Theory of Social Bonds (van Anders et al. 2011) we expected that T would increase after a challenging father-child interaction and decrease after a harmonious father-child interaction. However, this expectation was not confirmed in our study. These null-results could partly be due to the fact that with manipulations of human interactions, not all participants experience or respond to the manipulation as intended. Thus, despite designing the interaction episodes to be either challenging or harmonious, these may not have been challenging or harmonious enough for *all* fathers. We did not take into account whether, for example, *some* fathers found it easy to distract their child during the challenging interaction and whether *some* fathers found it challenging to play for 15 min in front of a camera during the harmonious interaction.

Another explanation for our null-results could be that fathers' T response in challenging and harmonious interactions are modulated by other factors. For example, when fathers witness their child in distress, their T reactivity might be modulated by the cognitive appraisals of the distress (Kuo et al. 2015; Zilioli and Bird 2017). More specifically, when fathers empathize with the distressed child, we might expect T to decrease, but when the distress aggravates them, T may increase. Other examples of moderating factors are fathers' affective state, their attachment style, or other hormones such as oxytocin. When these factors are not taken into consideration, the different T patterns stay unnoticed. Therefore, more attention should be given to moderating factors in future studies.

Testosterone and Paternal Sensitivity

In line with our expectations, we found evidence that the more T decreased during the harmonious father-child interaction, the more sensitive fathers were during this interaction. But contrary to our expectations, there was a trend that the more T increased during a challenging father-child interaction, the more sensitive fathers were during this interaction. Our results however, should be interpreted with caution because the associations between fathers' T reactivity and sensitivity within the same father-child interaction were only found in the second interaction during each visit, not in the first interaction. The absence of an association between changes in fathers' T levels and their sensitivity within the first interaction might be due to the effect of being observed. The awareness of being observed could evoke stress and thus have the unwanted effect of arousing fathers' sympathetic nervous system, especially at the start of the home visit (i.e., during the first interaction). This stress response might have influenced the association between fathers' sensitivity and T changes during the first interaction.

We found a trend that a higher T increase during the challenging interaction was related to more sensitive behavior of the father during that interaction. This finding suggests that increases in T are not necessarily associated with insensitive fathering. This appears to contradict the common notion that increases in T facilitate aggressive behaviors (Carré et al. 2011) and thus interfere with sensitive parenting. However, Bos (2017) explains that an increase in T can also have a protecting function that is beneficial for caregiving behavior. An increased T is associated with an increased social vigilance and sensitivity to facial expressions (Bos et al. 2012), which could make the father more perceptive of the child's signals. Also, brain regions associated with approach behaviors (Kuo et al. 2012) and parental responsiveness (Bos et al. 2010) are activated when T increases. A rapid T increase might motivate the father to protect and care for his distressed child (Bos 2017). Our results suggest that parenting quality is most optimal when fathers' T system reacts in the expected direction given the context of the father-child interaction, i.e., a T decrease during a harmonious interaction and a T increase during a challenging interaction.

We exploratively examined the temporal direction of the association between T and sensitivity. In both visits, fathers' T reactivity during the first father-child interaction predicted fathers' sensitivity during the second father-child interaction. Specifically, the more fathers' T levels *decreased* during the harmonious interaction, the more sensitive they were toward their child during the subsequent challenging interaction. And the more fathers' T levels *increased* during the challenging interaction, the more sensitive they were toward their child during the subsequent harmonious interaction. The results demonstrate that when fathers' T system reacts in the expected direction given the context of the father-child interaction (i.e., the more T increased during the challenging interaction, or decreased during the harmonious interaction), the more sensitive fathers were during that interaction as well as during a subsequent interaction. The reverse was not found, that is, sensitivity during the first father-child interaction did not predict changes in T levels during the second father-child interaction. Thus, the quality of the father-child interactions seems to depend, at least in part, on the T change during father-child interactions. This is in line with administration studies that point out that acute T increases are followed by behavioral changes (Zilioli and Bird 2017). Moreover, in their innovative study, Kuo et al. (2015) also found that T change in response to an interaction with their child modulated fathers' sensitivity during a subsequent father-child interaction. We build on the study of Kuo et al. (2015) by counterbalancing two different types of father-child interactions, i.e., during one visit father and child first interacted in the harmonious task and subsequently in the challenging task and during the other visit the order of these tasks was reversed. Moreover, we measured T and sensitivity during all father-child interactions. This enabled us to establish temporality and thus get a clearer impression of the direction of the association between T and sensitivity during different types of father-child interactions. However, the reversed direction cannot be ruled out, i.e., that fathers' characteristics and behaviors predict their T levels. It is difficult to determine causality in human research, because many aspects cannot be

controlled for and an association can partly be spurious. It remains possible that the effect of a third variable causes variation in both T and sensitivity, such as fathers' cognitive appraisals (Kuo et al. 2015; Zilioli and Bird 2017).

Strengths and Limitations

Our study's major strengths are the extensive T measurements and taking into account several aspects of fathers' parenting. The present study is also strengthened by the use of home observations. This ecologically-valid design enabled us to observe fathers' natural behaviors when interacting with their child. Another strength of our study is that we were able to test T reactivity specifically linked to the father-child interactions, by correcting for normal fluctuations of T levels during the afternoon. Moreover, by coding the amount of verbal and physical contact between father and child after each observation episode until right before the T measurements, we were able to account for the effect of unintended father-child contact on fathers' T levels. A final strength of our study was using an excellent methodology for measuring salivary T, i.e., LC-MS/MS (Higashi 2012).

Despite these strengths, the present study also has some limitations that could be addressed in future studies. A first limitation is our small non-representative sample. Our sample consisted predominantly of highly educated Caucasian fathers living in the Western region of the Netherlands. Moreover, a small sample size limits the power to detect associations between variables. This could partly explain why some findings just failed to reach significance. We have therefore chosen to interpret marginally significant results (with caution). Consequently, our findings require replication using larger and more diverse samples. Second, we did not use the same observational measures for examining sensitivity in the different father-child interactions. Fathers' sensitivity scores did not correlate between the different types of interactions. Although both measures were based on Mary Ainsworth's construct of sensitivity, the Erickson scales (Egeland et al. 1990) focus more on the cognitive side of sensitivity, while the Emotional Availability Scales (Biringen 2008) accentuate the affective side of sensitivity (Mesman and Emmen 2013). A third limitation is that we did not observe child behavior. Kuo et al. (2015) found that child distress was a predictor of fathers' sensitivity. Because behaviors of the child could confound the association between fathers' T levels and their sensitivity, future studies should take the child's behaviors into account.

Conclusion

To our knowledge, this is the first study that examined changes in fathers' T levels in relation to their sensitivity during different types of real-life father-child interactions. Our study indicates that fathers are most sensitive when their T levels increase during a challenging father-child interaction, and decrease during a harmonious father-child interaction. Moreover, changes in fathers' T levels seem to predict how fathers interact with their child, rather than the other way around. In sum, fathers' T reactivity seems to be beneficial for the quality of their parenting when T reacts in the expected direction given the context of the father-child interaction.

Acknowledgements We are grateful to Eric Vander Sypt for carrying out the hormonal analyses.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Alvergne, A., Faurie, C., & Raymond, M. (2009). Variation in testosterone levels and male reproductive effort: Insight from a polygynous human population. *Hormones and Behavior*, 56(5), 491–497. <https://doi.org/10.1016/j.yhbeh.2009.07.013>.
- Berg, S. J., & Wynne-Edwards, K. E. (2001). Changes in testosterone, cortisol, and estradiol levels in men becoming fathers. *Mayo Clinic Proceedings*, 76(6), 582–592. <https://doi.org/10.4065/76.6.582>.
- Berg, S. J., & Wynne-Edwards, K. E. (2002). Salivary hormone concentrations in mothers and fathers becoming parents are not correlated. *Hormones and Behavior*, 42(4), 424–436. <https://doi.org/10.1006/hbeh.2002.1841>.
- Biringen, Z. (2008). The emotional availability (EA) scales, 4th edition. Infancy/early childhood version (child age: 0–5 years). Retrieved from www.emotionalavailability.com.
- Booth, A., Granger, D. A., Mazur, A., & Kivlighan, K. T. (2006). Testosterone and social behavior. *Social Forces*, 85(1), 167–191. <https://doi.org/10.1353/sof.2006.0116>.
- Bos, P. A. (2017). The endocrinology of human caregiving and its intergenerational transmission. *Development and Psychopathology*, 29(3), 971–999. <https://doi.org/10.1017/S0954579416000973>.
- Bos, P. A., Hermans, E. J., Montoya, E. R., Ramsey, N. F., & van Honk, J. (2010). Testosterone administration modulates neural responses to crying infants in young females. *Psychoneuroendocrinology*, 35(1), 114–121. <https://doi.org/10.1016/j.psyneuen.2009.09.013>.
- Bos, P. A., Hermans, E. J., Ramsey, N. F., & van Honk, J. (2012). The neural mechanisms by which testosterone acts on interpersonal trust. *Neuroimage*, 61(3), 730–737. <https://doi.org/10.1016/j.neuroimage.2012.04.002>.
- Carré, J. M., McCormick, C. M., & Hariri, A. R. (2011). The social neuroendocrinology of human aggression. *Psychoneuroendocrinology*, 36(7), 935–944. <https://doi.org/10.1016/j.psyneuen.2011.02.001>.
- Edelstein, R. S., Chopik, W. J., Saxbe, D. E., Wardecker, B. M., Moors, A. C., & LaBelle, O. P. (2017). Prospective and dyadic associations between expectant parents' prenatal hormone changes and postpartum parenting outcomes. *Developmental Psychobiology*, 59(1), 77–90. <https://doi.org/10.1002/dev.21469>.
- Egeland, B., Erickson, M. F., Moon, J. C., Hiestler, M. K., & Korfmacher, J. (1990). *Revised Erickson Scales: 24 month tools coding manual. Project STEEP-revised 1990. From mother-child project scales 1978*. Minneapolis: University of Minnesota, Department of Psychology.
- Endendijk, J. J., Hallers-Haaboorn, E. T., Groeneveld, M. G., van Berckel, S. R., van der Pol, L. D., Bakermans-Kranenburg, M. J., et al. (2016). Diurnal testosterone variability is differentially associated with parenting quality in mothers and fathers. *Hormones and Behavior*, 80, 68–75. <https://doi.org/10.1016/j.yhbeh.2016.01.016>.
- Erickson, M. F., Sroufe, L. A., & Egeland, B. (1985). The relationship between quality of attachment and behavior problems in preschool in a high-risk sample. *Monographs of the Society for Research in Child Development*, 50(1–2), 147–166. <https://doi.org/10.2307/3333831>.
- Fleming, A. S., Corter, C., Stallings, J., & Steiner, M. (2002). Testosterone and prolactin are associated with emotional responses to infant cries in new fathers. *Hormones and Behavior*, 42(4), 399–413. <https://doi.org/10.1006/hbeh.2002.1840>.
- Gettler, L. T. (2016). Becoming DADS: Considering the role of cultural context and developmental plasticity for paternal socioendocrinology. *Current Anthropology*, 57(S13), S38–S51. <https://doi.org/10.1086/686149>.

- Gettler, L. T., McDade, T. W., & Kuzawa, C. W. (2011). Cortisol and testosterone in Filipino young adult men: Evidence for co-regulation of both hormones by fatherhood and relationship status. *American Journal of Human Biology*, 23(5), 609–620. <https://doi.org/10.1002/ajhb.21187>.
- Gettler, L. T., McKenna, J. J., McDade, T. W., Agustin, S. S., & Kuzawa, C. W. (2012). Does cosleeping contribute to lower testosterone levels in fathers? Evidence from the Philippines. *PLoS One*, 7(9), e41559. <https://doi.org/10.1371/journal.pone.0041559>.
- Gettler, L. T., McDade, T. W., Agustin, S. S., Feranil, A. B., & Kuzawa, C. W. (2015). Longitudinal perspectives on fathers' residence status, time allocation, and testosterone in the Philippines. *Adaptive Human Behavior and Physiology*, 1(2), 124–149. <https://doi.org/10.1007/s40750-014-0018-9>.
- Gray, P. B., Campbell, B. C., Marlowe, F. W., Lipson, S. F., & Ellison, P. T. (2004). Social variables predict between-subject but not day-to-day variation in the testosterone of US men. *Psychoneuroendocrinology*, 29(9), 1153–1162. <https://doi.org/10.1016/j.psyneuen.2004.01.008>.
- Hallers-Haalboom, E. T., Mesman, J., Groeneveld, M. G., Endendijk, J. J., van Berkel, S. R., van der Pol, L. D., et al. (2014). Mothers, fathers, sons and daughters: Parental sensitivity in families with two children. *Journal of Family Psychology*, 28(2), 138–147. <https://doi.org/10.1037/a0036004>.
- Higashi, T. (2012). Salivary hormone measurement using LC/MS/MS: Specific and patient-friendly tool for assessment of endocrine function. *Biological and Pharmaceutical Bulletin*, 35(9), 1401–1408. <https://doi.org/10.1248/bpb.b212009>.
- Kuo, P. X., Carp, J., Light, K. C., & Grewen, K. M. (2012). Neural responses to infants linked with behavioral interactions and testosterone in fathers. *Biological Psychology*, 91(2), 302–306. <https://doi.org/10.1016/j.biopsycho.2012.08.002>.
- Kuo, P. X., Saini, E. K., Thomason, E., Schultheiss, O. C., Gonzalez, R., & Volling, B. L. (2015). Individual variation in fathers' testosterone reactivity to infant distress predicts parenting behaviors with their 1-year-old infants. *Developmental Psychobiology*. <https://doi.org/10.1002/dev.21370>.
- Kuzawa, C. W., Gettler, L. T., Muller, M. N., McDade, T. W., & Feranil, A. B. (2009). Fatherhood, pairbonding and testosterone in the Philippines. *Hormones and Behavior*, 56(4), 429–435. <https://doi.org/10.1016/j.yhbeh.2009.07.010>.
- Lamb, M. E. (2000). The history of research on father involvement: An overview. *Marriage and Family Review*, 29(2–3), 23–42. https://doi.org/10.1300/J002v29n02_03.
- Lamb, M. E. (2010). *The role of the father in child development* (5th ed.). Hoboken: John Wiley & Sons.
- Lawson, D. W., Nunez-de la Mora, A., Cooper, G. D., Prentice, A. M., Moore, S. E., & Sear, R. (2017). Marital status and sleeping arrangements predict salivary testosterone levels in rural Gambian men. *Adaptive Human Behavior and Physiology*, 3(3), 221–240. <https://doi.org/10.1007/s40750-017-0066-z>.
- Matsui, F., Koh, E., Yamamoto, K., Sugimoto, K., Sin, H.-S., Maeda, Y., et al. (2009). Liquid chromatography-tandem mass spectrometry (LC-MS/MS) assay for simultaneous measurement of salivary testosterone and cortisol in healthy men for utilization in the diagnosis of late-onset hypogonadism in males. *Endocrine Journal*, 56(9), 1083–1093. <https://doi.org/10.1507/endocrj.K09E-186>.
- Mesman, J., & Emmen, R. A. (2013). Mary Ainsworth's legacy: A systematic review of observational instruments measuring parental sensitivity. *Attachment & Human Development*, 15(5–6), 485–506. <https://doi.org/10.1080/14616734.2013.820900>.
- Perini, T., Ditzen, S., Fischbacher, U., & Ehlert (2012). Testosterone and relationship quality across the transition to fatherhood. *Biological Psychology*, 90(3), 186–191, doi: <https://doi.org/10.1016/j.biopsycho.2012.03.004>.
- Pleck, J. H. (2010). Paternal involvement: Revised conceptualization and theoretical linkages with child outcomes. In M. E. Lamb (Ed.), *The role of the father in child development* (pp. 58–93). Hoboken: Wiley.
- Pruessner, J. C., Kirschbaum, C., Meinlschmid, G., & Hellhammer, D. H. (2003). Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. *Psychoneuroendocrinology*, 28(7), 916–931. [https://doi.org/10.1016/S0306-4530\(02\)00108-7](https://doi.org/10.1016/S0306-4530(02)00108-7).
- Roney, J. R., Lukaszewski, A. W., & Simmons, Z. L. (2007). Rapid endocrine responses of young men to social interactions with young women. *Hormones and Behavior*, 52(3), 326–333. <https://doi.org/10.1016/j.yhbeh.2007.05.008>.
- Storey, A. E., Walsh, C. J., Quinton, R. L., & Wynne-Edwards, K. E. (2000). Hormonal correlates of paternal responsiveness in new and expectant fathers. *Evolution and Human Behavior*, 21(2), 79–95. [https://doi.org/10.1016/S1090-5138\(99\)00042-2](https://doi.org/10.1016/S1090-5138(99)00042-2).
- Storey, A. E., Noseworthy, D. E., Delahunty, K. M., Halfyard, S. J., & McKay, D. W. (2011). The effects of social context on the hormonal and behavioral responsiveness of human fathers. *Hormones and Behavior*, 60(4), 353–361. <https://doi.org/10.1016/j.yhbeh.2011.07.001>.
- Tabachnick, B. G., & Fidell, L. S. (2012). *Using multivariate statistics* (6th ed.). New York: Harper Collins.

- van Anders, S. M., Goldey, K. L., & Kuo, P. X. (2011). The steroid/peptide theory of social bonds: Integrating testosterone and peptide responses for classifying social behavioral contexts. *Psychoneuroendocrinology*, 36(9), 1265–1275. <https://doi.org/10.1016/j.psyneuen.2011.06.001>.
- van Anders, S. M., Tolman, R. M., & Volling, B. L. (2012). Baby cries and nurturance affect testosterone in men. *Hormones and Behavior*, 61(1), 31–36. <https://doi.org/10.1016/j.yhbeh.2011.09.012>.
- van Anders, S. M., Goldey, K. L., & Bell, S. N. (2014). Measurement of testosterone in human sexuality research: Methodological considerations. *Archives of Sexual Behavior*, 43(2), 231–250. <https://doi.org/10.1007/s10508-013-0123-z>.
- Weisman, O., Zagoory-Sharon, O., & Feldman, R. (2014). Oxytocin administration, salivary testosterone, and father-infant social behavior. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 49, 47–52. <https://doi.org/10.1016/j.pnpbp.2013.11.006>.
- Wingfield, J. C., Hegner, R. E., Dufty, A. M., & Ball, G. F. (1990). The "challenge hypothesis": Theoretical implications for patterns of testosterone secretion, mating systems, and breeding strategies. *American Naturalist*, 136(6), 829–846. <https://doi.org/10.1086/285134>.
- Zilioli, S., & Bird, B. M. (2017). Functional significance of men's testosterone reactivity to social stimuli. *Frontiers in Neuroendocrinology*, 47, 1–18. <https://doi.org/10.1016/j.yfme.2017.06.002>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Else E. de Vries¹ · Lotte D. van der Pol² · Harriet J. Vermeer¹ · Marleen G. Groeneveld² · Tom Fiers³ · Judi Mesman¹

Lotte D. van der Pol
l.d.van.der.pol@fsw.leidenuniv.nl

Harriet J. Vermeer
vermeer@fsw.leidenuniv.nl

Marleen G. Groeneveld
mgroeneveld@fsw.leidenuniv.nl

Tom Fiers
Tom.Fiers@uzgent.be

Judi Mesman
mesmanj@fsw.leidenuniv.nl

¹ Education and Child Studies, Leiden University, P.O. Box 9555, 2300 RB Leiden, The Netherlands

² Faculty of Governance and Global Affairs, Leiden University, P.O. Box 13228, 2501 EE Den Haag, The Netherlands

³ Clinical Chemistry, Ghent University Hospital, De Pintelaan 185, 9000 Ghent, Belgium